Welcome: 9:30-9:40 (EST)

Session 1: 9:40-10:40 (EST)
Chair: Yulia Gel

Xiaojing Ye – Georgia State University
Inference and Prediction via Neural Mean-Field Dynamics on Networks

We consider a novel learning framework based on neural mean-field dynamics for inference and estimation problems of diffusion on large heterogeneous networks. Our new framework is derived from the Mori-Zwanzig formalism to obtain an exact evolution of the node infection probabilities, which renders a delay differential equation with memory integral approximated by learnable time convolution operators, resulting in a highly structured and interpretable RNN. Directly using cascade data, our framework can jointly learn the structure of the diffusion network and the evolution of individual node infection probabilities, which are cornerstone to important downstream applications such as influence maximization. Connections between parameter learning and optimal control are also established. Empirical study shows that our approach is versatile and robust to variations of the underlying diffusion network models, and significantly outperform existing approaches in accuracy and efficiency on both synthetic and real-world test data.

Austin Benson – Cornell University
Spectral embeddings for evolving networks

Algorithms for deriving features from graph or network data often assume that the computations are over a static graph. However, in many cases, we gather network data over time or from various sources; in this sense, the networks are evolving. In this talk, we will cover our recent research analyzing how spectral embeddings, a key feature for graphs, changes as a network evolves. Our results are centered on a new random graph model for evolving networks as well as novel perturbation theory and distributed algorithms for computing spectral embeddings.

Yao Xie – Georgia Institute of Technology
Uncertainty quantification and sequential change detection for Hawkes networks

Multivariate Hawkes processes are commonly used to model streaming networked event data in various applications, including neural science, social networks, seismic data, crime data, and epidemiology. Much progress in estimating such models has been made in statistics and machine learning; however, it remains a challenge to perform uncertainty quantification and extract
reliable inference from complex datasets, especially considering general interaction patterns. This is essential for statistical tasks such as casual inference (where the underlying directed graph implies Granger causality) and change detection. Aiming towards this, we develop statistical inference tools for finding sequential confidence intervals and detecting changes, drawing ideas from concentration inequalities of continuous-time martingales and optimization. We compare our method to the previously derived asymptotic Hawkes process confidence interval and demonstrate our method's strengths in an application to neuronal connectivity reconstruction.

Session 2: 11:00-12:00 (EST)
Chair: Yao Xie

George Michailidis – University of Florida

*Online Detection of Changes in High-Dimensional Graphical Models*

High dimensional piecewise stationary graphical models represent a versatile class for modelling time varying networks arising in diverse application areas, including biology, economics, and social sciences. We introduce a novel scalable online algorithm for detecting an unknown number of abrupt changes in the inverse covariance matrix of sparse Gaussian graphical models with small delay. The proposed algorithm is based upon monitoring the conditional log-likelihood of all nodes in the network and can be extended to a large class of continuous and discrete graphical models. We also investigate asymptotic properties of our procedure under certain mild regularity conditions on the graph size, sparsity level, number of samples, and pre- and post-changes in the topology of the network. Numerical works on both synthetic and real data illustrate the good performance of the proposed methodology both in terms of computational and statistical efficiency across numerous experimental settings.

Yajun Mei – Georgia Institute of Technology

*Adaptive Sequential Change-Point Detection with Sampling Control*

In the context of real-time monitoring multiple data streams, it is assumed that there are M processes in a system and at some unknown time, an occurring event impacts some unknown local processes in that sense of changing the distributions of observations from those affected local process. In this talk, we present our latest results of such a problem under the sampling control, in which one is able to take observations from some local processes at each time step. First, we develop an efficient change-point detection algorithm with sampling control that turns out to be second-order asymptotically optimal in the scenario when exactly one local process is affected by the change. That is, with the sampling rate that is only $1/M$ of the full data scenario, our proposed algorithm has the same performance up to second-order as the optimal procedure under the full data scenario. Second, we discuss how to combine our proposed algorithms with multi-armed bandit algorithms to develop efficient methods in more complicated scenarios such as when multiple local data streams might be affected or when the post-change distributions are composite due to multiple failure models. Finally, we will also briefly our results on the hot-spots detection for COVID-19 in the state of Georgia. These research are based on our collaborative project funded in September 2018 entitled “Adaptive and Rapid Spatial-Temporal Threat
Detection Over Networks” with Prof. Sarah Holte from Fred Hutchinson Cancer Research Center in Seattle and Prof. Hao Yan from Arizona State University.

Yulia Gel – University of Texas at Dallas
*Anomaly Detection on Complex Networks with Topological Data Analysis*

We introduce a new nonparametric method for anomaly detection on graphs using the emerging tools of topological data analysis. Our method is based on accounting for changes in a topological and geometric structure of dynamic networks under the framework of persistent homology. In particular, we describe shapes of complex networks via analysis of clique communities and then track fluctuations of the resulting topological signatures over time. We validate the proposed approach in application to unilayer and multilayer networks from social sciences, telecommunication and blockchain.

**Session 3 (Challenge): 1:00-1:30 (EST)**
Chair: John Greer

Doug Mercer – Penn State Applied Research Lab
*ATD2020 Challenge Results*

In this year’s ATD challenge problem, participants detected anomalous observations in sparsely sampled traffic flow data. In this talk, we’ll provide an overview of the challenge problem statement, dataset, and evaluation methodology. We’ll conclude by presenting the challenge results and congratulating the winning team and runners-up to this year’s ATD challenge.

Wes Whiting – University of California, Irvine
*An Ensemble of Random Forests for Anomaly Detection*

The ATD2020 challenge posed the following problem: given very incomplete traffic data from road sensors, how can we tell which readings are outliers? This talk will discuss the approach we used in our winning submission, incorporating geographic and correlational relationships between sensors.

**Session 4: 2:00-3:00 (EST)**
Chair: Abel Rodriguez

Shane Bookhultz – Virginia Tech
*Modeling and Understanding Polarization in News Media using novel Dynamic Topic Models*

Polarization in social media and news media reflects the polarization in society, especially in the last few years. Social science researchers have established that this polarization often threatens civil society through acts of violence. In this project, we employ topic modeling to build an automatic threat detection tool, one that analyzes news media and provides a barometer for polarization. Specifically, we apply topic modeling techniques to detect essential topics in the
news cycle and model the news variation in terms of topic persistence, new topic emergence, and topic absorption. We quantify the influence of these temporally varying topics by using a high-dimensional multilevel version of a system of Dynamic Linear Models. With the combined implementation of topic models and time varying models, the proposed method accounts for both time and topic dependence. Using a variant of sentiment analysis, we create a polarization measure that tracks with the polarization in different news topics over time, and use this measure to demonstrate underlying associations between topic polarity and known highly polarizing events.

Georgios Fellouris – University of Illinois, Urbana-Champaign
*Sequential Identification of a Dependence Structure*

We will consider a setup where observations are collected sequentially in multiple sources and the problem is to identify the subset of dependent sources, as quickly as possible, subject to various error constraints. We will establish the optimal average sample size, to a first-order asymptotic approximation, for every possible dependence structure under consideration. Moreover, we will propose a scheme that is computationally simple in that the number of recursions required per sampling instance is quadratic in the number of sources. We will compute the asymptotic relative efficiency of the proposed scheme under various dependence structures and illustrate its actual performance with simulation studies.

Daniel Sanz-Alonso – University of Chicago
*Graphical approximations of Matern Gaussian fields: theory and applications*

In this talk I will introduce graph representations of stochastic partial differential equations as a way to approximate Matern Gaussian field. Approximation error guarantees will be established building on and generalizing the theory of spectral convergence of graph Laplacians. Graph representations allow inference and sampling with linear algebra methods for sparse matrices, thus reducing the computational cost of Gaussian field approaches. In addition, they bridge and unify several models in Bayesian inverse problems, spatial statistics and graph-based machine learning. We demonstrate through examples in these three disciplines that the unity revealed by graph representations facilitates the exchange of ideas across them. This is joint work with Ruiyi Yang.

**Posters: 3:30-5:00 (EST)**

<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Affiliation</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Michael Baron</td>
<td>American University</td>
<td>Sequential Analysis of Epidemic Threats</td>
</tr>
<tr>
<td>2</td>
<td>Autumn Chadwick</td>
<td>Rowan University</td>
<td>Stock and Flow Inventory Modeling Using Markov Chains</td>
</tr>
<tr>
<td>3</td>
<td>Anamitra Chaudhuri</td>
<td>University of Illinois, Urbana Champaign</td>
<td>Quickest Detection and Isolation of a Dependent Pair</td>
</tr>
<tr>
<td></td>
<td>Name</td>
<td>Affiliation</td>
<td>Title</td>
</tr>
<tr>
<td>----</td>
<td>----------------------</td>
<td>------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>4</td>
<td>Taewon Cho</td>
<td>Virginia Tech</td>
<td>Hybrid Projection Method for Large-scale Inverse Problems with Mean Estimation in Hierarchical Gaussian Priors</td>
</tr>
<tr>
<td>5</td>
<td>Harris Hardiman-Mostow</td>
<td>Tufts University</td>
<td>Anomaly Detection in Sparsely Sampled Traffic Flow</td>
</tr>
<tr>
<td>6</td>
<td>Charles Harrison</td>
<td>University of Central Florida</td>
<td>Efficient Bayesian Clustering for Social Network Data</td>
</tr>
<tr>
<td>7</td>
<td>Qing He</td>
<td>University of Central Florida</td>
<td>Bayesian Clustering for Enron Emailing Network Data</td>
</tr>
<tr>
<td>8</td>
<td>Shushan He</td>
<td>Georgia State University</td>
<td>Network Diffusions via Neural Mean-Field Dynamics</td>
</tr>
<tr>
<td>9</td>
<td>Abigail Hickok</td>
<td>UCLA</td>
<td>A Bounded-Confidence Model on Hypergraphs</td>
</tr>
<tr>
<td>10</td>
<td>Shen-Shyang Ho</td>
<td>Rowan University</td>
<td>Detecting Anomaly in Evolving Graph using Martingale</td>
</tr>
<tr>
<td>11</td>
<td>Jiuyun Hu</td>
<td>Arizona State University</td>
<td>Adaptive Partially-Observed Sequential Change Point Detection for Covid-19 hotspots detection</td>
</tr>
<tr>
<td>12</td>
<td>Jiahua Jiang</td>
<td>ShanghaiTech University</td>
<td>Hybrid Projection Methods for Large-scale Inverse Problems with Solution Decomposition</td>
</tr>
<tr>
<td>13</td>
<td>Latifur Khan</td>
<td>University of Texas Dallas</td>
<td>A Unified Spatio-Temporal Graph Neural Network for Time Series Forecasting</td>
</tr>
<tr>
<td>14</td>
<td>Michael Kirby</td>
<td>Colorado State University</td>
<td>Self-Organizing Mappings on the Flag Manifold</td>
</tr>
<tr>
<td>15</td>
<td>Dustin Mixon</td>
<td>The Ohio State University</td>
<td>Sketching semidefinite programs for faster clustering</td>
</tr>
<tr>
<td>16</td>
<td>Kaiyan Peng</td>
<td>UCLA</td>
<td>Network embedding for multiplex online social networks</td>
</tr>
<tr>
<td>17</td>
<td>Yuyang Shi</td>
<td>Georgia Institute of Technology</td>
<td>Kernel Smoothing based Anomaly Detection Algorithm for ATD2020 Challenge</td>
</tr>
<tr>
<td>18</td>
<td>Martin Short</td>
<td>Georgia Institute of Technology</td>
<td>A Self-limiting Hawkes Process: Interpretation, Estimation, and Use in Crime Modeling</td>
</tr>
<tr>
<td>19</td>
<td>Seyyid Emre Sofuoglu</td>
<td>Michigan State University</td>
<td>Graph Regularized Low Rank Plus Temporally Smooth Sparse Tensor Decomposition for Anomaly Detection in Spatio-Temporal Data</td>
</tr>
<tr>
<td>20</td>
<td>Charisopoulos Vasileios</td>
<td>Cornell University</td>
<td>Communication-efficient distributed eigenspace estimation</td>
</tr>
<tr>
<td>21</td>
<td>Han Xiao</td>
<td>Rutgers</td>
<td>CoD: Conformal Outlier Detection</td>
</tr>
<tr>
<td>22</td>
<td>Hao Yan</td>
<td>Arizona State University</td>
<td>Adaptive Partially-Observed Sequential Change Point Detection with Multiple Failure Modes</td>
</tr>
<tr>
<td>23</td>
<td>Zhou Yu</td>
<td>University of Illinois at Chicago</td>
<td>Smoothing Spline ANOVA Analysis for Traffic Data</td>
</tr>
<tr>
<td>24</td>
<td>Minghe Zhang</td>
<td>Georgia Institute of Technology</td>
<td>Unsupervised spatio-temporal anomaly detection with missing data</td>
</tr>
<tr>
<td></td>
<td>Xinyu Zhao</td>
<td>Arizona State University</td>
<td>Hierarchical Tree-based Sequential Event Prediction with Application in the Aviation Accident Report</td>
</tr>
<tr>
<td>---</td>
<td>----------------</td>
<td>--------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>26</td>
<td>Lin Zhu</td>
<td>Texas A&amp;M University</td>
<td>Point process models for terrorism attacks</td>
</tr>
</tbody>
</table>
Tuesday, 11/10/2020

Session 5: 9:30-10:30 (EST)
Chair: James Murphy

Ping Ma – University of Georgia
Subsampling in Streaming Time Series

Advances in data acquisition technology pose challenges in analyzing large volumes of streaming data. Subsampling is a natural yet powerful tool for analyzing such data sets due to their competent estimation accuracy and low computational cost. Unfortunately, Subsampling methods and their statistical properties for streaming data, especially streaming time series data, are not well studied in the literature. In this talk, I will present an online leverage-based sequential subsampling algorithm for streaming time series data. Simulation studies and real data examples are presented to evaluate the empirical performance of the proposed method.

Jie Yang – University of Illinois at Chicago
Identifying zero-inflated distributions

Count data with a large portion of zeros arise naturally in many scientific disciplines, including security, insurance, health care, microbiome, and more. When conducting one-sample Kolmogorov-Smirnov (KS) test for count data, the estimated p-value is biased due to plugging in sample estimates of unknown parameters. As a consequence, the result of a KS test could be too conservative. In the newly developed R package iZID for zero-inflated count data, we use bootstrapped Monte Carlo estimates to overcome the bias issue in estimating p-values, as well as bootstrapped likelihood ratio tests for zero-inflated model selection. Our new package also provides miscellaneous functions to simulate zero-inflated count data and calculate maximum likelihood estimates of unknown parameters. Compared with other R packages available so far, our package covers more types of zero-inflated distributions and provides adjusted p-value estimates after incorporating the influence of unknown model parameters.

Haomin Zhou – Georgia Institute of Technology
Learning stochastic behavior from aggregated data

Learning nonlinear dynamics from aggregate data is a challenging problem since the full trajectory of each individual is not available, namely, the individual observed at one time point may not be observed at next time point, or the identity of individual is unavailable due to technical limitations, experimental cost and/or privacy issues. This is in sharp contrast to learning dynamics with trajectory data, on which the majority of existing methods are based. In this paper, we present a complementary strategy designed for aggregate data. Our method uses weak form of Fokker Planck Equation (FPE) to describe density evolution of data in a sampling form, which is combined with Wasserstein generative adversarial network (WGAN) in training process. In such
a sample-based framework we are able to study nonlinear dynamics from data without solving
the partial differential equation (PDE). The model can also handle high dimensional cases with
the help of neural networks. We demonstrate our approach in the context of a series of synthetic
and real-world data sets.

**Session 6: 11:00-12:00 (EST)**
Chair: Ping Ma

Gilad Lerman – University of Minnesota

*Novelty Detection via Robust Variational Autoencoding*

We propose a new method for novelty detection that can tolerate high corruption of the training
points, whereas previous works assumed either no or very low corruption. Our method trains a
robust variational autoencoder (VAE), which aims to generate a model for the uncorrupted
training points. To gain robustness to high corruption, we incorporate the following four changes
to the common VAE: 1. Extracting crucial features of the latent code by a carefully designed
dimension reduction component for distributions; 2. Modeling the latent distribution as a
mixture of Gaussian low-rank inliers and full-rank outliers, where the testing only uses the inlier
model; 3. Applying the Wasserstein-1 metric for regularization, instead of the Kullback-Leibler
(KL) divergence; and 4. Using a least absolute deviation error for reconstruction. We establish
both robustness to outliers and suitability to low-rank modeling of the Wasserstein metric as
opposed to the KL divergence. We illustrate state-of-the-art results on standard benchmarks for
novelty detection.

James Murphy – Tufts University

*Multiscale diffusion metrics for weakly-supervised learning*

Methods for unsupervised and semi-supervised learning on large, sparsely connected networks
are developed. We propose metrics and related sampling schemes that exploit multiscale
structure in the networks to provably capture latent cluster hierarchies. These methods enjoy
robust theoretical guarantees and fast numerical schemes that scale. Applications to link and
function prediction for incomplete and noisy protein-protein interaction networks are
emphasized.

Eric Weber – Iowa State University

*Efficient and Stable Algorithms for Non-Euclidean Regression in Discrete Geometries*

We consider the problem of linear regression when the data is distributed across a network. We
propose an algorithm for solving the linear regression problem and demonstrate that it is stable.
Our results apply to several different network topologies. We will present an overview of our
methods, as well as discuss other results obtained by our team of researchers.
Session 7: 1:00-2:00 (EST)
Chair: Subhadeep Paul

Alexander Vladimirsky – Cornell University
Quantifying & managing uncertainty in piecewise-deterministic Markov processes

In piecewise-deterministic Markov processes (PDMP) the state of a finite-dimensional system evolves dynamically, but the evolutive equation may change randomly as a result of discrete switches. A running cost is integrated along the corresponding piecewise-deterministic trajectory up to the termination to produce the cumulative cost of the process. We address three natural questions related to uncertainty in cumulative cost of PDMP models: (1) how to compute the Cumulative Distribution Function (CDF) when the switching rates are fully known; (2) how to accurately bound the CDF when the switching rates are uncertain; and (3) assuming the PDMP is controlled, how to select a control to optimize that CDF. In all three cases, our approach requires posing a (weakly-coupled) system of suitable hyperbolic partial differential equations, which are then solved numerically on an augmented state space. We illustrate our method using simple examples of trajectory planning under uncertainty. Joint work with E. Cartee, A. Farah, A. Nellis, and J. van Hook.

Sun Yuekai – University of Michigan
There is no trade-off: enforcing fairness can improve accuracy

One of the main barriers to the broader adoption of algorithmic fairness in machine learning is the trade-off between fairness and performance of ML models: many practitioners are unwilling to sacrifice the performance of their ML model for fairness. In this paper, we show that this trade-off may not be necessary. If the algorithmic biases in an ML model are due to sampling biases in the training data, then enforcing algorithmic fairness may improve the performance of the ML model on unbiased test data. We study conditions under which enforcing algorithmic fairness helps practitioners learn the Bayes decision rule for (unbiased) test data from biased training data. We also demonstrate the practical implications of our method.

Mason Porter – University of California, Los Angeles
Bounded-Confidence Models on Networks

There are a large variety of models of opinion dynamics. In so-called "bounded-confidence models", entities have continuous-valued opinions, and they compromise with each other, by adjusting their opinion states, when they interact. By examining a bounded-confidence model on a network, one can examine how social connectivity patterns affect such opinion dynamics. In this talk, I will introduce bounded-confidence models, discuss the cascading of opinions that can occur when entities influence each other in such models, and examine bounded-confidence opinion dynamics on networks with both dyadic and polyadic interactions.
Session 8: 2:30-3:50 (EST)
Chair: Sun Yuekai

Hsin-Hsiung Huang – University of Central Florida
*Unsupervised AIS trajectory reconstruction, affine-invariant and matrix-variate clustering, and robust discriminant analysis.*

In this talk, I will present three of my recent research studies supported by the NSF ATD grant. Our research team has developed a Bayesian clustering framework which includes robust discriminant analysis, affine-transformation invariant clustering, and identifying the groups of subjects in networks, proposed the unsupervised method of vessel movement trajectory prediction, and investigated its properties. The developed unsupervised clustering algorithms which have high accuracy with low computational complexity provide real-time solutions to detect dynamic pattern changes. Among them, the robust discriminant analysis and affine-transformation invariant clustering models have been accepted for publication supported by the NSF ATD grant. From these two papers, our research team has developed several new research projects of robust clustering for matrix-variate data and variable selection.

Subhadeep Paul – The Ohio State University
*Testing for network small world property*

Researchers have long observed that the “small-world” property is ubiquitous for complex networks obtained from diverse disciplines. However, we find two shortcomings in the currently popular definition and detection methods rendering the concept somewhat less powerful. First, the popular definition combines high transitivity or clustering with a low average path length in a somewhat ad-hoc fashion, which confounds the two. We find that most networks detected as “small world” with the existing methodology are merely a manifestation of their high transitivity. Second, the detection methods lack a formal statistical test, and the comparison is typically performed against simplistic random graph models. We propose two innovations to make the property more informative. First, we define the property as a statistical test between a suitable null and superimposed alternative models. The test performs parametric bootstrap with several null models representing a varied degree of structure in the network. Second, we decouple the properties of high transitivity and low average path length as separate criteria to test for. In addition to the bootstrap tests, we also propose an asymptotic test under the Erdos-Renyi null model for which we provide theoretical guarantees on the asymptotic level and power. Applying the proposed methods on a large number of network datasets, we uncover new insights about their small-world property that would not have been possible otherwise.

Mikyoung Jun – University of Houston
*Multivariate Spatio-temporal Hawkes process models for terrorist attacks*

We develop flexible spatio-temporal Hawkes process models to analyze patterns of terrorism. Proposed models are multivariate providing flexible structure for "cross triggering", nonstationary accounting for complex spatial and temporal structure of the point patterns, and
non-separable in space-time domain. To demonstrate the utility of models' flexibility, we analyze two samples of real-world data. One is terrorist attack patterns by Taliban in Afghanistan, and the other is bivariate terrorism patterns by Boko Haram and Fulani Extremists in Nigeria. Our analyses demonstrate that our model dramatically outperforms standard Hawkes process models of both the Afghanistan and Nigerian terror data. It not only bests widely used alternatives in overall model fit, but reveals spatio-temporal patterns that are, by construction, masked in these models. This is a joint work with Scott Cook.

Rebecca Willett – University of Chicago

_Detection and Description of Change in Visual Streams_

I will describe a framework for the analysis of changes in visual streams: ordered sequences of images, possibly separated by significant time gaps. We propose a new approach to incorporating unlabeled data into training to generate natural language descriptions of change. We also develop a framework for estimating the time of change in a visual stream. We use learned representations for change evidence and consistency of perceived change and combine these in a regularized graph-cut based change detector. Experimental evaluation on visual stream datasets shows that representation learning driven by natural language descriptions significantly improves change detection accuracy, compared to methods that do not rely on language. This is joint work with Davis Gilton, Ruotian Luo, and Greg Shakhnarovich.